Correlation Between System Usability, Mental Workload and their impact on Human Machine Interaction

MOHAMMAD DELWAR HOSSIN, Frankfurt University of Applied Sciences, Germany
OMME SALMA, Frankfurt University of Applied Sciences, Germany
NIMA, Frankfurt University of Applied Sciences, Germany
MONIKA, Frankfurt University of Applied Sciences, Germany
SANDRA, Frankfurt University of Applied Sciences, Germany

Currently, our environments are increasingly being populated with smart devices. To provide such convenient smart devices we need to integrate wide functionality in a single interface which can increase mental workload [4]. This interface complexity increases mental workload which is one of the important factors for any system. But Some researchers suggest that there is no relation between usability and mental workload. On the other hand, Some researchers' studies show that there is a correlation. In this paper, we investigated this controversial finding by collecting SUS and Nasa TLX data from 24 research papers and then analyzed these data and checked their relationship using linear regression and the Pearson correlation algorithm. Finally, summarized our findings which suggest that there is a strong negative correlation between usability and mental workload which impact Human-machine Interaction.

CCS Concepts: • **Human-centered computing** → **Usability testing**; • **General and reference** → **Surveys and overviews**.

Additional Key Words and Phrases: Usability, SUS, Mental workload, NASA TLX, correlation

ACM Reference Format:
Mohammad Delwar Hossin, Omme Salma, Nima, Monika, and Sandra. 2022. Correlation Between System Usability, Mental Workload and their impact on Human Machine Interaction. *Proc. ACM Meas. Anal. Comput. Syst.* 37, 4, Article 111 (August 2022), 10 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

In the modern era, interface complexity is increasing for integrating wide functionality in any single system to make it a robust or multifunctional system. This interface complexity increases mental workload which is one of the important factors for any system that results in errors in safety and task performance. Therefore, Usability and Mental workload are important factors in any technological device, system or web interface. Generally, Usability refers to how usable an application or system seems to be, whereas mental workload means how demanding a task appears to be for the user. These two factors play important role in the success or failure of system, application or technological device [9]. Systems with good usability are more likely to use by many people as these systems are easy to use and tend to a lower mental workload on any user.

Authors' addresses: Mohammad Delwar Hossin, Frankfurt University of Applied Sciences, Frankfurt am Main, Germany, mohammad.hossin@stud.fra-uas.de; Omme Salma, Frankfurt University of Applied Sciences, Frankfurt am Main, Germany, omme.salma@stud.fra-uas.de; Nima, Frankfurt University of Applied Sciences, Frankfurt am Main, Germany; Monika, Frankfurt University of Applied Sciences, Frankfurt am Main, Germany; Sandra, Frankfurt University of Applied Sciences, Frankfurt am Main, Germany.

2476-1249/2022/8-ART111 $15.00
https://doi.org/XXXXXXX.XXXXXXX

Over the past few decades importance of usability and mental workload has become increasingly emphasized in the system development phase. It is a necessary condition for a website or system. Poor usability can increase the mental workload as users have to give more effort or mental effort to use the system. According to some researchers [9] ,[3] studies show that usability and mental workload are two non-overlapping constructs in any interactive system. There is no correlation between usability and Mental workload. Some researchers' studies show that there is a correlation between these two factors. A study by Ariza et al [7] shows that Poor usability can increase the mental workload experienced by clinicians and cause fatigue which increases error rates. Another study by Christina M. Kokin et al.[10] suggests that mental workload strongly correlated with the perceived usability of a system. So, there is a mismatch among these researchers' findings. Our main research aim is to work on this controversial issue.

A controversial issue arises in the relationship between system usability and mental workload in different research papers. A key aspect of user interaction is usability which is a quality attribute that assesses how easy user interfaces are to use. It refers to the quality of a user's experience when interacting with systems, including websites, software, devices, or applications. The success of any system mainly depends upon the usability of the user's opinion which can only be achieved by integrating a user-centered design approach early in the design process. For example, a website with good usability is one that provides a user-friendly design and easy-to-use interface which can increase user satisfaction [10] . Human Mental Workload is fundamental for exploring the interaction of people with the technological device. It is the amount of mental work necessary for a person to complete a task over a given period of time that reflects the number of mental resources required to perform a set of concurrent tasks. It depends on the characteristics of the interface and the type of task, and it may affect both user performance and usability.

Today's human activities are increasingly cognitively driven [9]. Therefore, the assessment of mental workload and usability has become an important aspect of designing effective systems that reduce human errors. There are many subjective procedures for measuring mental workload and usability. QUIS, ASQ SMEQ and SUS (System Usability Score) approaches are used for measuring usability, and NASA TLX, WP used for measuring subjective mental workload. Among them, SUS and NASA Tlx are the most widely used methods. SUS is a post-test questionnaire developed by John Brooke in 1986 and has become recognized as an industry standard with references in over 1300 articles and publications. It measures the usability of products, including hardware, software, mobile devices, websites and applications. It contains 10-item questions answered using 5 options Likert scale from strongly disagree to strongly agree [13]. NASA Task Load Index or NASA TLX is also a post-test questionnaire subjective workload assessment technique that contains 6-item questions answered using 5 options Likert scale from Low to high and Good to Poor. It collects data using survey and relies on a multidimensional construct to derive an overall workload score based on a weighted average of ratings on six subscales: mental , physical, temporal, performance, effort, and frustration level [6]

The aim of this research is to investigate the relationship between usability and mental workload. We considered SUS and NASA TLX as subjective measurement tools for measuring usability and mental workload. We conducted our whole literature review task on the basis of the following research questions "Is there a relationship between System Usability and Human mental workload which impact on Human-Machine Interaction? ".To explore this relationship, We collected 54 sample data of SUS and corresponding NASA TLX from different research papers. Following that, we applied the linear regression algorithm and Pearson's correlation to these data to investigate their relationship as our hypothesis was Usability and Mental workload are two correlated constructs in Human-Computer Interaction.

This paper is organized as follows. At first, we provide research background, an Overview of Usability and Mental workload, and their assessment techniques such as NASA TLX and SUS. In the next sections, we present some related work and then explain our research methodology. In sections 3 and 4, we analyze collected data and proved our research hypothesis. In the last section, we provide a summary of this paper.

2 BACKGROUND AND MOTIVATION

2.1 Problem and Motivation

Usability is a necessary condition for a website or system. Poor usability can increase the mental workload as users have to give more effort or mental effort to use the system. According to some researchers Luca Longo and Longo et.al [9],[3] studies show that usability and mental workload are two non-overlapping constructs in any interactive system which means there is no correlation between usability and Mental workload. Some researchers' studies show that there is a correlation between these two factors. A study by Ariza et al. [7] shows that "Poor usability can increase the mental workload experienced by clinicians and cause fatigue, increase error rates". Another study by Christina M. Kokin et al. [4] suggests that mental workload and operator control were strongly correlated with the perceived usability of the system. So, there is a mismatch arises among these researchers' findings. Our main research aim is to work on this controversial issue. Therefore, we want to investigate the correlation between usability and mental workload.

2.2 Objective

The purpose of this research is to find out the relationship between usability and mental workload using subjective NASA Tlx and SUS questionnaire scores which are collected from different research papers.

2.3 Reasearch Question and Hypothesis

We conducted our whole literature review task on the basis of the following research questions.

Is there a relationship between System Usability and Human mental workload which impacts on Human-Machine Interaction?

Our research hypothesis is: "Usability and Mental workload are two correlated constructs in Human Interaction system".

2.4 Usability

Usability is a quality attribute that assesses how easy user interfaces are to use. It refers to the quality of a user's experience when interacting with systems, including websites, software, devices, or applications. For example, a website with good usability is one that allows a user-friendly design and easy-to-use interface which can increase user satisfaction. There are a number of post-study usability testing methods available for assessment. These are QUIS, SEQ, SMEQ, CSUQ, ASQ, and SUS. Almost all of these questionnaires have some distinct features, including reliability assessment, validity, and sensitivity, making them valuable tools for usability testers. Among them, SUS and CSUQ are the most popular post-test questionnaires that are widely used and readily available [9],[25].

2.4.1 SUS. It is a post-test questionnaire developed by John Brooke in 1986 and has become recognized as an industry standard with references in over 1300 articles and publications. It measures the usability of products and services, including hardware, software, mobile devices, websites and applications. It collects data using survey that contains 10-item questions answered using 5 options Likert scale from strongly disagree to strongly agree. A SUS score is presented as a single number ranging from 0 to 100 where 68 is considered to be an average score. It actually measures how user-friendly people think the system is, but it does not measure how easy it is to actually use the system.

2.5 Mental Workload

Human Mental Workload is an important design concept and it is fundamental for exploring the interaction of people with technological devices. Mental workload is the amount of mental work necessary for a person to complete a task over a given period of time. It depends on the characteristics of the interface and the nature of

the task, and it may affect both user performance and usability[1]. There are a number of methods available to measure mental workloads such as WP, and NASA TLX. All assessment techniques should possess the following properties such as ease of implementation, low intrusiveness, and reliability. Among them, Nasa TLX is the most popular as it fulfills all of the assessment features [6].

2.5.1 NASA TLX. The NASA TLX is a six-item questionnaire developed in 1980 by Sandra Hart. It measures workload which means the amount of effort has to pay to use any human interaction system. It relies on a multidimensional construct to provide workload scores based on a weighted average of ratings on six subscales such as mental demand, physical demand, temporal demand, performance, effort, and frustration. It is a tool for measuring and conducting a subjective mental workload (MWL) assessment which calculates the workload of a participant while they are performing a task [1].

3 RELATED WORK

In the following, we provide an overview of the previous work that is related to the measurement of subjective usability, and mental workload using NASA Tlx and SUS approaches. This includes research on the relationship between NASA TLX and SUS score in the context of human-machine interaction to investigate how SUS score impacts mental workload. A poor SUS score increases mental workload and a higher SUS score reduces mental workload. Finally, we discuss how SUS score impacts on mental workload score. We collected data from different research papers where authors use this approach in different application scenarios.

In developing a system, the importance of system usability has become a prominent factor to achieve better human performance. Good usability is a core component of the success or failure of a system. A study by Christina M. Kokin et. al [9] showed that both mental workloads had significant effects on perceived usability. Also, the perceptions of mental workload are strongly correlated with the perceived usability of the system. Another study by Ariza et. al [3] suggests that Clinical information systems in the National Health Service do not need to conform to any explicit usability requirements. They found that "poor usability can increase the mental workload experienced by clinicians and cause fatigue, increase error rates and impact the overall patient safety. The mental workload can be used as a measure of usability" [3].

Further research of the Nasa Tlx and SUS questionnaire by Leanne M et. al [7] on Brain measurement for usability testing and adaptive interfaces shows that the mental workload required to complete a task is the sum of task difficulty and complexity of the user interface. They conducted their study using an experimental protocol where users had to complete some predefined task on the computer. Another similar study conducted by Longo et.al [10] show that "The tasks performed on Wikipedia, increments in required mental workload correspond to decrements in usability perception and vice-versa, underlying an inverted correlation".Their finding suggests that there is an inverse relationship between mental workload and usability.

Krause et.al [14] conducted a user study on Subjective Ratings in an Ergonomic Engineering Process. They used a self-reporting approach NASA Tlx and SUS to assess the web interface and their result indicate that a weak negative correlation between NASA-TLX and SUS exist. In contrast, a study conducted by Luca Longo [12] on perceived usability and mental workload where participants had to do some typical predefined tasks. After that Nasa Tlx and SUS questionnaires were given to provide feedback. Their result shows that usability and mental workload are two independent constructs. These findings completely differs from previously [9],[3], [7],[10],[14] mentioned results.

Previous work shows that there is a controversial issue arises as some researchers [9], [3], [10] said usability and mental workload are dependent constructs where some authers [12], [13] results showed that these two constructs are completely independent. Previously mentioned authors conducted their study by designing a user study and a controversy arise in their findings. In this sense, our work differs from the above-mentioned studies.

Instead of conducting a user study, we collected data from different research papers and find their correlation to investigate these controversies.

4 METHOD

Our research purpose is to investigate the relationship between mental workload and usability. For this investigation, we need to find those papers where authors used Nasa TLX and SUS Questionnaire. Following that, we screen and filter those papers by analyzing our research need as we need both SUS and its corresponding NASA Tlx score. In this section, we will explain our data acquisition, extraction, analysis strategy, and an overview of the research methodology. The following figures represent the overview of our research methodology.

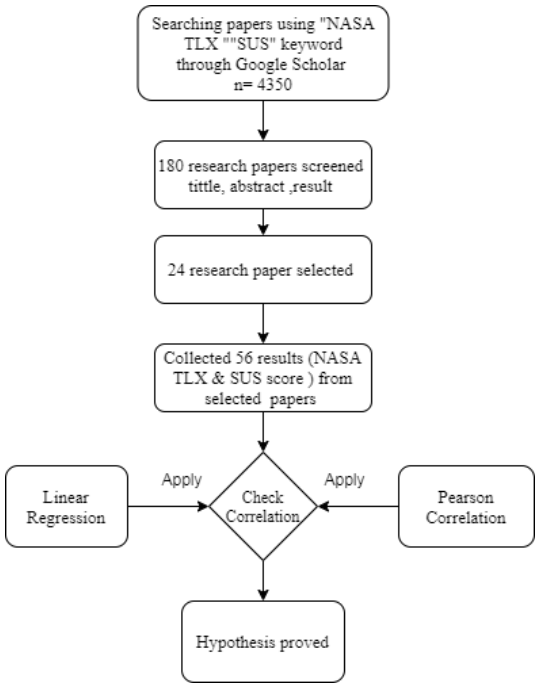

Fig. 1. An Overview of research methodology for correlation checking between SUS and Nasa Tlx

4.1 Data Acquisition and Extraction

There are a number of methods available to measure the system's usability and mental workload. We have used two subjective measurement approaches NASA Tlx and SUS questionnaire. We collected data from different research papers where authors uses these two method. As Usability and mental workload measurement is a wide field of research therefore we started our study by searching the target words ("Usability ", "Mental workload") on Google Scholar and IEEE digital library. After performing the search 4350 research papers have found on Google Scholar, some are related to the web usability inspection methods, and some are related to the mental workload but in most of the research papers authors used different subjective measurement approaches such as QUIS, ASQ, and SMEQ). We then re-performed our search keywords as "Nasa Tlx", and "SUS" as our research task is to find out the relationship between these two subjective measurement approaches. Then, we found the expected papers from where we collected NASA TLX and SUS score data. We extracted our research data from these research

papers by reading the Abstract and Result sections. After screening and reading 180 relevant papers, we found 24 papers [2, 5, 8, 11, 15–17, 19–24, 26**? ?**] where different researchers used NASA TLX and SUS methods.

4.2 Data Analysis

Before started screening papers, we first defined our searching criteria as " NASA TLX" and "SUS". In total 180 papers were included whereas 22 papers related to our research need. We read the result section properly and checked SUS score and its corresponding NASA TLX score. We found 24 papers where authors conducted different user studies in multiple fields and then collected SUS and NASA TlX scores. We have collected 54 data samples from these 24 papers and analyzed these samples by applying Linear Regression and Pearson's correlation algorithm to investigate our research hypothesis.

5 RESULT

5.1 Collected Dataset

In total, we read 180 papers from the Google Scholar database and selected 24 papers which is relevant to our research need. We collected and analyzed 54 usability and corresponding mental workload data from these papers. Figure 2 contains 54 SUS and NASA TLX scores from 24 research papers. From an initial analysis of Figure 3, it seems clear that there is a Negative correlation between SUS and NASA TLX. This is statistically proved in table 1 by the Pearson correlation computed over the full dataset. Pearson was chosen for exploring linear correlation.

	sus	nasa						
1	55.83	75.81	19	80.00	30.00	36	45.00	34.00
2	47.78	67.57	20	66.50	41.50	37	80.00	29.10
3	80.42	38.83	21	90.00	16.00	38	76.00	28.07
4	44.17	75.22	22	88.00	18.00	39	71.25	29.31
5	44.44	73.03	23	86.00	17.00	40	37.29	56.37
6	86.11	35.00	24	85.00	19.00	41	55.94	51.08
7	78.03	21.03	25	79.40	23.70	42	71.46	44.95
8	34.03	52.04	26	76.70	25.00	43	80.00	40.00
9	50.00	49.80	27	68.70	35.30	44	54.00	45.00
10	59.00	54.00	28	80.80	32.20	45	56.00	50.00
11	68.00	51.00	29	80.20	33.30	46	38.00	64.00
12	80.00	42.00	30	58.00	38.80	47	50.00	55.00
13	68.25	43.01	31	77.16	25.06	48	60.00	40.00
14	91.83	22.00	32	61.68	68.09	49	69.40	18.80
15	51.70	44.00	33	71.47	23.34	50	66.10	21.80
16	90.40	25.00	34	76.25	22.08	51	74.00	30.00
17	73.30	32.00	35	77.92	22.67	52	80.00	27.00
18	91.83	23.00	36	45.00	34.00	53	56.20	58.50
			37	80.00	29.10	54	59.00	49.20

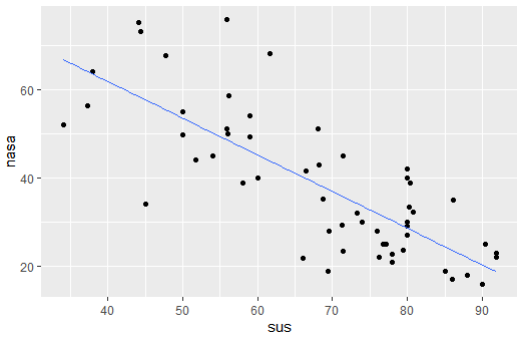

Fig. 2. Collected SUS and NASA TLX scores from different researchers' findings (right) and relationship between SUS and NASA TLX score (right)

5.2 Correlation result: SUS and NASA TLX

5.2.1 Linear Regression. To test the hypothesis, SUS value and NASA-TLX value were collected. we applied a linear regression algorithm to collected data. To establish a linear regression algorithm on sample data, SUS is considered to be independent and NASA TlX is the dependent variable. Here, we can see in figure 3, we found a significant relationship between SUS and NASA Tlx score $F(1;53) = 81.62$; where $p < .001$, Also in Coefficient correlation showed significant negative correlation between NASA TLX and SUS where $r = -0.779$ and $p < :001$. Generally, the correlation scale is range from [-1,1] where above 0.5 are regarded as very high positive, within [0.1, 0.3] small, and within [0.3 0.5] as medium correlation. Similarly, [-0.5, -1] is considered to be highly negative

correlation. Since we found the coefficient value as (r= -0.779) which proved that NASA TLX negatively correlated with SUS. This suggests that when the proportion of SUS score is decreased mental workload is increased. Conversely, if usability is increased then the mental workload is decreased.

Model Summary - nasa

Model	R	R^2	Adjusted R^2	RMSE
H_0	0.000	0.000	0.000	16.192
H_1	0.779	0.606	0.599	10.255

ANOVA

Model		Sum of Squares	df	Mean Square	F	p
H_1	Regression	8584.237	1	8584.237	81.621	< .001
	Residual	5574.098	53	105.172		
	Total	14158.336	54			

Note. The intercept model is omitted, as no meaningful information can be shown.

Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p
H_0	(Intercept)	38.463	2.183		17.616	< .001
H_1	(Intercept)	94.962	6.405		14.827	< .001
	sus	-0.829	0.092	-0.779	-9.034	< .001

Fig. 3. Statistical Summary results of Correlation between SUS and NASA TLX

5.2.2 Pearson correlation. Figure 4 shows Pearson Correlation performed on the whole dataset To test our hypothesis, further analysis was done using Pearson correlation which measures the strength of the linear relationship between two variables. Correlation scale is range from [-1,1] where -1 negative correlation, 0 belongs to no correlation, and +1 denotes a positive correlation. Here, we can see that we found a statistically significant negative relationship between SUS and NASA Tlx score,r = -0.779; where p <.001. Therefore, it strongly proved our hypothesis that usability and mental workload are two correlated constructs. Finally, statistical results proved that there is an inverse relationship between SUS and NASA TLX score.

To further strengthen the hypothesis result, a density plot of the correlation between Mental workload and usability scores has been performed in figure 4. This density plot also confirmed our hypothesis that there is a negative correlation between SUS and Nasa Tlx.

Pearson's Correlation

Pearson's Correlations

			n	Pearson's r	p	Lower 95% CI	Upper 95% CI
sus	-	nasa	55	-0.779***	< .001	-0.865	-0.647

* p < .05, ** p < .01, *** p < .001

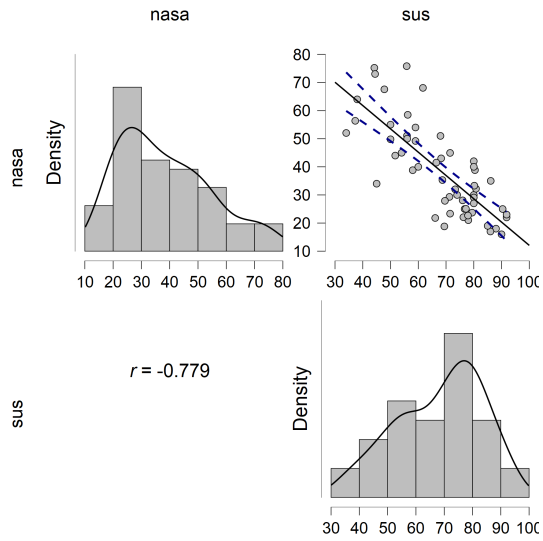

Fig. 4. Statistical Summary results of person Correlation (Ist) and Density plot of SUS and NASA TLX score

5.3 Summary of findings

In summary, from meta-analysis evidence, the hypotheses can be accepted.

"Usability and Mental workload are two correlated constructs in the human-computer Interaction" (as measured with the selected technique (SUS, NASA-TLX). They capture a negative correlation. This has been tested by linear regression, person correlation also density plot analysis which confirmed that the two constructs are strongly correlated with each other. Linear regression results showed a significantly strong correlation between these constructs where $p < .001$. The obtained Pearson coefficients also suggest that there was a negative correlation between usability (SUS scale) and mental workload (NASA-TLX) which was ($r = -0.779$). Similarly, by analyzing the density plot, it was also confirmed that there is an inverse relation between SUS and NASA TLX that is Good usability score helps to reduce mental workload, on the other hand, a poor usability score tends to increase mental workload.

6 CONCLUSION

The main purpose of this literature review paper was to investigate the correlation between the perception of usability and mental workload by collecting data from the different research papers. In the first two sections, we discussed usability and workload assessment tools and also provided an overview of our research methodology. Two well-known subjective assessment techniques such as NASA TLX, SUS have been selected to collect usability and mental workload data. We collected SUS and Corresponding NASA TLX data from 23 research papers and then analyzed these data and checked their relationship using linear regression and Pearson correlation algorithm. Following that, summarized our findings which suggest that there is a strong negative correlation between usability and mental workload which impact Human-machine Interaction. Finally, our hypothesis has been checked as true which means usability and Mental workload are two correlated constructs.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support and motivation of Prof. Dr. Valentin Schwind, Faculty of Computer Science and Engineering, Frankfurt University of Applied Sciences. Following his inspiration and continuous guideline, it is made possible to complete this paper.

REFERENCES

[1] Adem Akbiyik, Khaled Hassanein, and Milena Head. 2015. Perceived maintenance load of High Maintenance Information Systems: Conceptualization and development. (2015).

[2] David E Applebury, Eric J Robinson, Jeffrey A Gold, Jeffrey D Davis, and David Zonies. 2021. Pilot testing of simulation in the evaluation of a novel, rapidly deployable Electronic Health Record for use in disaster intensive care. *Disaster Med. Public Health Prep.* (Oct. 2021), 1–5.

[3] Ferran Ariza, Dipak Kalra, and Henry Ww Potts. 2015. How do clinical information systems affect the cognitive demands of general practitioners?: Usability study with a focus on cognitive workload. *J. Innov. Health Inform.* 22, 4 (Nov. 2015), 379–390.

[4] Adam Ben-Hanania, Joshua Goldberg, and Jessica R Cauchard. 2018. Multi-user control for domestic robots with natural interfaces. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia* (Cairo Egypt). ACM, New York, NY, USA.

[5] Caroline Brendle, Laura Schütz, Javier Esteban, Sandro M Krieg, Ulrich Eck, and Nassir Navab. 2020. Can a hand-held navigation device reduce cognitive load? A user-centered approach evaluated by 18 surgeons. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Springer International Publishing, Cham, 399–408.

[6] Alex Cao, Keshav K Chintamani, Abhilash K Pandya, and R Darin Ellis. 2009. NASA TLX: software for assessing subjective mental workload. *Behav. Res. Methods* 41, 1 (Feb. 2009), 113–117.

[7] Leanne M Hirshfield, Erin Treacy Solovey, Audrey Girouard, James Kebinger, Robert J K Jacob, Angelo Sassaroli, and Sergio Fantini. 2009. Brain measurement for usability testing and adaptive interfaces. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09* (Boston, MA, USA). ACM Press, New York, New York, USA.

[8] Gretchen Hultman, Reed McEwan, Serguei Pakhomov, Elizabeth Lindemann, Steven Skube, and Genevieve B Melton. 2017. Usability evaluation of NLP-PIER: A clinical document search engine for researchers. *Stud. Health Technol. Inform.* 245 (2017), 1269.

[9] Christina M Kokini, Sangwon Lee, Richard J Koubek, and Seung Ki Moon. 2012. Considering context: The role of mental workload and operator control in users' perceptions of usability. *Int. J. Hum. Comput. Interact.* 28, 9 (Sept. 2012), 543–559.

[10] Michael Krause and Klaus Bengler. 2013. Subjective ratings in an ergonomic engineering process using the example of an in-vehicle information system. In *Human-Computer Interaction. Applications and Services*. Springer Berlin Heidelberg, Berlin, Heidelberg, 596–605.

[11] Michael Krause and Klaus Bengler. 2013. Subjective ratings in an ergonomic engineering process using the example of an in-vehicle information system. In *Human-Computer Interaction. Applications and Services*. Springer Berlin Heidelberg, Berlin, Heidelberg, 596–605.

[12] Luca Longo. 2017. Subjective Usability, Mental Workload Assessments and Their Impact on Objective Human Performance. In *Human-Computer Interaction - INTERACT 2017*, Regina Bernhaupt, Girish Dalvi, Anirudha Joshi, Devanuj K. Balkrishan, Jacki O'Neill, and Marco Winckler (Eds.). Springer International Publishing, Cham, 202–223.

[13] Luca Longo. 2018. Experienced mental workload, perception of usability, their interaction and impact on task performance. *PLoS One* 13, 8 (Aug. 2018), e0199661.

[14] Luca Longo. 2018. Experienced mental workload, perception of usability, their interaction and impact on task performance. *PLoS One* 13, 8 (Aug. 2018), e0199661.

[15] Nicolas S Marjanovic, Agathe De Simone, Guillaume Jegou, and Erwan L'Her. 2017. A new global and comprehensive model for ICU ventilator performances evaluation. *Ann. Intensive Care* 7, 1 (Dec. 2017), 68.

[16] Jackie McSorley and Beth Blickensderfer. 2019. Usability analysis of convective SIGMETs. (2019).

[17] Jackie McSorley and Beth Blickensderfer. 2019. Usability analysis of convective SIGMETs. (2019).

[18]]L Emanuele ; Pasqualotto, A Simonetta, ; Veronica, S ; Olivetti Federici, and M Belardinelli. [n. d.]. Toward a Usability Evaluation of BCIs. *International Journal* ([n. d.]).

[19] Emanuele Pasqualotto, A Simonetta, Veronica Gnisci, S Federici, and M Olivetti Belardinelli. 2011. Toward a usability evaluation of BCIs. *Int. J. Bioelectromagn.* 13 (2011), 121.

[20] Emanuele Pasqualotto, A Simonetta, Veronica Gnisci, S Federici, and M Olivetti Belardinelli. 2011. Toward a usability evaluation of BCIs. *Int. J. Bioelectromagn.* 13 (2011), 121.

[21] R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

[22] Patricia Viertler, Stephan Schlögl, Reinhard Mayer, Matthias Janetschek, and Juliana Pattermann. 2021. Show me the universe! Perceived usability and task load of an AR mobile-app in secondary school learning. In *Communications in Computer and Information Science*. Springer International Publishing, Cham, 40–50.

[23] Patricia Viertler, Stephan Schlögl, Reinhard Mayer, Matthias Janetschek, and Juliana Pattermann. 2021. Show me the universe! Perceived usability and task load of an AR mobile-app in secondary school learning. In *Communications in Computer and Information Science*. Springer International Publishing, Cham, 40–50.

[24] Stefan Werrlich, Austino Daniel, Alexandra Ginger, Phuc-Anh Nguyen, and Gunther Notni. 2018. Comparing HMD-based and paper-based training. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Munich, Germany). IEEE.

[25] P Whittington, H Dogan, and K Phalp. 2015. Evaluating the usability of an Automated Transport and Retrieval System. In *2015 International Conference on Pervasive and Embedded Computing and Communication Systems (PECCS)*. IEEE, 49–66.

[26] P Whittington, H Dogan, and K Phalp. 2015. Evaluating the usability of an automated transport and retrieval system. In *2015 International Conference on Pervasive and Embedded Computing and Communication Systems (PECCS)*. 49–66.